A modified thymine for the synthesis of site-specific thymine-guanine DNA interstrand crosslinks

Alzeer, Jawad ; Schärer, Orlando D

Abstract: DNA interstrand crosslinks (ICLs) are highly cytotoxic lesions formed by a variety of important anti-tumor agents. Despite the clinical importance of ICLs, the mechanisms by which these lesions are repaired in mammalian cells have so far remained elusive. One of the obstacles in the study of ICL repair has been the limited availability of suitable methods for the synthesis of defined site-specific ICLs. We report here the synthesis of a site-specific ICL containing an ethylene-bridged G-T base pair based on the incorporation of a crosslink precursor containing a selectively reactive group on one strand using solid-phase synthesis. 3-(2-chloroethyl)thymidine was incorporated into oligonucleotides and underwent ICL formation upon annealing to a complementary strand by reacting with the base opposite to the modified T residue. A strong preference for ICL formation with a G residue opposite the reactive T was observed. Detailed characterization of the reaction product revealed that the alkylation reaction occurred with the O-6 group of G and a mechanism accounting for this preference is proposed. These G-T crosslinks introduced here will be useful for studies of ICL repair.

DOI: https://doi.org/10.1093/nar/gkl587

4458–4466 *Nucleic Acids Research, 2006, Vol. 34, No. 16* *Published online 31 August 2006*
doi:10.1093/nar/gkl587

A modified thymine for the synthesis of site-specific thymine-guanine DNA interstrand crosslinks

Jawad Alzeer[1] and Orlando D. Schärer[1,2,*]

[1]Institute of Molecular Cancer Research, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland and [2]Department of Pharmacological Sciences, Chemistry Graduate Building 619, Stony Brook University, Stony Brook, NY11794-3400

Received June 22, 2006; Revised July 26, 2006; Accepted July 28, 2006

ABSTRACT

DNA interstrand crosslinks (ICLs) are highly cytotoxic lesions formed by a variety of important anti-tumor agents. Despite the clinical importance of ICLs, the mechanisms by which these lesions are repaired in mammalian cells have so far remained elusive. One of the obstacles in the study of ICL repair has been the limited availability of suitable methods for the synthesis of defined site-specific ICLs. We report here the synthesis of a site-specific ICL containing an ethylene-bridged G-T base pair based on the incorporation of a crosslink precursor containing a selectively reactive group on one strand using solid-phase synthesis. 3-(2-chloroethyl)-thymidine was incorporated into oligonucleotides and underwent ICL formation upon annealing to a complementary strand by reacting with the base opposite to the modified T residue. A strong preference for ICL formation with a G residue opposite the reactive T was observed. Detailed characterization of the reaction product revealed that the alkylation reaction occurred with the O-6 group of G and a mechanism accounting for this preference is proposed. These G-T crosslinks introduced here will be useful for studies of ICL repair.

INTRODUCTION

DNA interstrand crosslinks (ICLs) are among the most cytotoxic lesions known and are produced as a result of a reaction of cellular DNA with compounds containing bifunctional electrophilic moieties. The extraordinary cytotoxicity of ICLs is due to the fact that they block all cellular processes that involve the separation of dsDNA such as transcription or replication (1). ICL-forming bifunctional compounds are frequently used in cancer chemotherapy, but are also generated endogenously as metabolic side products (2). To counteract the threats of ICLs, cells have developed mechanisms to repair these lesions, and the failure to repair such lesions is associated with inherited disorders such as Fanconi Anemia (3,4). In contrast, the repair of ICLs in tumor cells leads to resistance against anti-tumor agents (5,6). The understanding of the mechanisms of ICL repair should, therefore, not only significantly contribute to our understanding of the molecular basis of human disease, but also to the mechanisms of resistance of tumor cells to chemotherapeutic agents. This knowledge should thus eventually facilitate the development of new targets for cancer therapy. Although significant progress has been made in studying various DNA repair pathways (7,8), the mechanisms underlying the repair of ICLs remain poorly understood (1,2,5,9). Apart from the inherent complexity of ICL repair, in which two damaged strands need to be repaired, the difficulty of preparing well-defined and stable ICL-containing oligonucleotides as substrates for biochemical and cell biological studies presents an additional significant hurdle in enabling the investigation of ICL repair pathways. Traditionally ICLs have been prepared by treating dsDNA with bifunctional alkylating agents, but this approach offers little or no sequence specificity, produces complex mixtures of mono adducts, intra- and interstrand crosslinks and very low overall yield of the desired ICL (usually <5%) (10–14).

De novo chemical synthesis provides an alternative for the generation of site-specific ICLs. One approach has relied on the cross-linking of mono nucleosides outside of DNA and the incorporation of crosslinked dimers into DNA using solid-phase synthesis after appropriate functionalization and protection (15–20). This approach has been useful for the preparation of defined short oligonucleotides, but is somewhat limited in scope due to the need of up to three different protecting groups on the phosphoramidite of the dimer. A practically more straightforward approach to the synthesis of ICLs consists in the incorporation of crosslink precursor in complementary single-stranded oligonucleotides and installation of the crosslink using a suitable coupling reaction after annealing of the two oligonucleotides. This approach has been used with precursors in one (21–23) or both strands of DNA (24–26).

As part of our program to study the repair of ICLs formed by bifunctional alkylating agents used in cancer

*To whom correspondence should be addressed. Tel: +1 631 632-77545; Fax: +1 631 632-7546; Email: orlando@pharm.stonybrook.edu

chemotherapy, we were interested in developing synthetic methods for the preparation of ICLs containing an ethylene bridge between two Watson–Crick paired bases to mimic the lesion formed by the chloro ethyl nitroso ureas (CENUs) (14,27). We describe a new approach for the preparation of ethylene bridge cross-linked DNA using a thymine containing a chloroethyl group at the N-3 position in a single-stranded oligonucleotide as a precursor. When paired to a complementary oligonucleotide with the modified T opposite to G, efficient ICL formation was observed. Interestingly, we found that of the two possible alkylation products with N-1 or O-6 of guanine, only reaction with the O-6 position was observed. Our studies thus provide a new approach to the synthesis of defined DNA interstrand crosslinks.

MATERIALS AND METHODS

All solvents and reagents were purchased from Sigma-Aldrich. Anhydrous solvents were prepared by distillation under an argon atmosphere in the presence of appropriate drying agents. Flash chromatography was performed using Silica gel 60 (220–430 mesh) as the adsorbent. Thin layer chromatography (TLC) was performed using precoated TLC plates (Merck, Silica gel 60 F_{254}, 0.25 mm).

^{1}H-NMR and ^{13}C-NMR spectra were recorded on a Bruker 300 MHz NMR spectrometer operating at room temperature. Chemical shifts are reported in parts per million (δ) downfield from tetramethylsilane. Spin multiplicities are described as s (singlet), bs (broad singlet), d (doublet), dd (doublet of doublets), t (triplet) and m (multiplet). Mass spectra were recorded by ESI-MS (Finnigan MAT TSQ-700) and oligonucleotides were measured on Q-TOF Ultima Mass Spectrometer (Micromass, UK) with a nano-source in negative mode. The spectra were obtained by deconvolution of the multiply charged ions using the MaxEnt1 program (Micromass, UK). HPLC was performed on a JASCO system equipped with a BGB Analytic column: Nucleosil 100 Å, C18, 5 mm 250 × 4.0 mm^2. The C18-SepPak cartridges were obtained from Micropore. Snake venom phosphodiesterase I (SVPD) was purchased from Worthington Biochemical; calf intestine phosphatase, from New England Biolabs.

N^3-(2-Chloroethyl-)-5′-O-dimethoxytrityl-3′-O-(tert-butyldimethylsilyl)thymidine (**2**). At room temperature, Cs_2CO_3 (1.33 g, 4.09 mmol) was added to a solution of 5′-*O*-dimethoxytrityl-3′-*O*-(*tert*-butyldimethylsilyl)thymidine **1** (28) (899 mg, 1.36 mmol) in DMF (8 ml). The mixture was stirred at room temperature for 20 min and then bromochloroethane (338 µl, 4.09 mmol) was added. The reaction mixture was stirred at room temperature for 17 h and treated with a saturated solution of NH_4Cl, extracted with ether, washed with brine and H_2O and dried over Na_2SO_4. Evaporation and silica gel column chromatography (Hexane/EtOAc 5:1) gave **2** (850 mg, 86%) as a white foam. R_f (Hexane/EtOAc 3:1) 0.67; ^{1}H NMR (300 MHz, $(CD_3)_2SO$): δ 7.66 (d, 1H, J = 1.1 Hz), 7.41 (dd, 2H, J = 7.2 Hz), 7.28 (m, 7H), 6.82 (d, 4H, J = 8.77 Hz), 6.36 (t, 1H, J = 6.26 Hz), 4.52 (m, 1H), 4.31 (dt, 2H, J = 1.7, 6.9 Hz), 3.98 (dd, 1H, J = 2.73, 5.85 Hz), 3.77 (s, 6H), 3.71 (t, 2H, J = 6.8 Hz), 3.48 (dd, 1H, J = 2.76, 10.66 Hz), 3.20 (dd, 1H, J = 2.85, 10.64 Hz), 2.34 (m, 1H), 2.20 (m, 1H), 1.52 (s, 3H), 0.84 (s, 9H), 0.03 (s, 3H), 0.00 (s, 3H); ^{13}C NMR (75 MHz, $(CD_3)_2SO$): δ 163.12, 158.65, 150.62, 144.13, 135.32, 133.90, 129.95, 128.04, 127.86, 127.01, 113.17, 109.99, 86.70, 86.69, 85.51, 71.94, 62.78, 55.11, 42.06, 41.51, 39.93, 25.60, 17.81, 12.50, −0.47, −0.49; ESI MS (M + Na^+): 743.3 (calcd 742.99).

N^3-(2-Chloroethyl-)-5′-O-dimethoxytritylthymidine (**3**). At room temperature, tetrabutylammonium fluoride (164 mg, 0.52 mmol) was added to a solution of N^3-(2-Chloroethyl-)-5′-*O*-dimethoxytrityl-3′-*O*-(*tert*-butyldimethylsilyl)thymidine **2** (250 mg, 0.34 mmol) in THF (3 mL). The mixture was stirred at room temperature for 1.5 h, treated with a saturated solution of NH_4Cl, extracted with EtOAc, washed with brine and H_2O and dried over Na_2SO_4. Evaporation and silica gel column chromatography (hexane/EtOAc 3:1) gave **3** (188 mg, 89%) as a colorless oil. R_f (Hexane/EtOAc 3:1) 0.30; ^{1}H NMR (300 MHz, $(CD_3)_2SO$): δ 7.58 (s, 1H), 7.41 (dd, 2H, J = 7.2 Hz), 7.33–7.23 (m, 7H), 6.88 (d, 4H, J = 8.7 Hz), 6.24 (t, 1H, J = 6.8 Hz), 5.33 (br s, OH), 4.33 (m, 1H), 4.13 (t, 2H, J = 6.8 Hz), 3.90 (m, 1H), 3.73 (s, 6H), 3.71 (m, 2H), 3.21 (m, 2H), 2.32–2.15 (m, 2H), 1.52 (s, 3H); ^{13}C NMR (75 MHz, $(CD_3)_2SO$): δ 162.34, 158.14, 158.03, 150.12, 144.54, 135.29, 135.15, 134.57, 129.77, 129.59, 127.75, 127.54, 126.65, 113.23, 113.11, 108.56, 85.74, 85.51, 84.73, 70.22, 63.53, 54.91, 41.53, 40.20, 39.67, 12.68; ESI MS (M + Na^+): 629.1 (calcd 629.2).

N^3-(2-Chloroethyl-)-5′-O-dimethoxytrityl-3′-O-(2-Cyanoethyl-N,N-diisopropylphosphor-amidite) thymidine (**4**). At room temperature, diisopropylethylamine (240 ml, 1.4 mmol) was added to a solution of N^3-(2-chloroethlyl-)-5′-*O*-dimethoxytritylthymidine **3** (322 mg, 0.53 mmol) in CH_2Cl_2 (3 ml), followed by chloro(2-cyanoethoxy)(*N*,*N*-diisopropylamino)phosphine (190 µl, 0.84 mmol). After 15 min, the reaction mixture was filtered through a short silica gel column (hexane/EtOAc 2:1), and fractions containing the desired product were combined and evaporated to dryness to afford **4** (mixture of diastereomers) as a white foam (389 mg, 90%). R_f (Hexane/EtOAc 3:1) 0.55; ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.62 (s, 1/2H), 7.58 (s, 1/2H), 7.37 (d, 2H, J = 8.8 Hz), 7.25 (m, 7H), 6.81 (d, 4H, J = 8.2 Hz), 6.39 (m, 1H), 4.63 (br. s, 1H), 4.27 (t, 2H, J = 6.68 Hz), 4.14 (d, 1H, J = 12.17 Hz), 3.76 (s, 6H), 3.70 (m, 3H), 3.54 (m, 4H), 3.30 (m, 1H), 2.30–2.60 (m, 4H), 1.44 (s, 3H), 1.15 (s, 9H, J = 5.9), 1.05 (d, 3H, J = 6.4 Hz); ^{31}P NMR {^{1}H} (300 MHz, $CDCl_3$): δ 149.65, 149.16; ESI MS (M^+): 806.5 (calcd 806.32).

N^3-(2-Iodoethyl-)-5′-O-dimethoxytrityl-3′-O-(tert-butyldimethylsilyl)thymidine (**11**) (20). At room temperature, NaI (83 mg, 0.56 mmol) was added to a solution of N^3-(2-chloroethyl-)-5′-*O*-dimethoxytrityl-3′-*O*-(*tert*-butyldimethylsilyl)thymidine **2** (213 mg, 0.29 mmol) in acetonitrile (3 ml). The reaction mixture was heated at 70°C for 18 h, cooled to room temperature, diluted with EtOAc, washed with sodium thiosulfate (10%, 3 × 5 ml), dried over Na_2SO_4 and evaporated to dryness to afford **11** (207 mg, 86%) as a light yellow foam. R_f (Hexane/EtOAc 3:1) 0.69. Analytical data were in accordance with previous reports for **11** (20).

1-{N^1-[3′,5′-O-(bis(tert-butyldimethylsilyl))-2-N-((dimethylamino)methylene)-2′-deoxyguanosinyl]}-2-{N^3-[5′-O-(dimethoxytrityl)-3′-O-(tert-butyldimethylsilyl)-thymidinyl]} ethane (13). At room temperature, Ag_2O (147 mg,

0.64 mmol) was added to a solution of N^3-(2-iodoethyl-)-5′-O-dimethoxytrityl-3′-O-(*tert*-butyldimethylsilyl)thymidine **11** (346 mg, 0.42 mmol) and 3′,5′-O-(bis(*tert*-butyldimethylsilyl))-2-N-((dimethylamino)methylene)-2′-deoxyguanosine **12** (228 mg, 0.42 mmol) in DMF (5 mL). The reaction mixture was kept at room temperature for 48 h, treated with a saturated solution of NH_4Cl, extracted with ether, washed with brine and H_2O and dried over Na_2SO_4. Evaporation and silica gel column chromatography (CH_2Cl_2:MeOH 20:1) gave **13** (250 mg, 48%) as an oil. R_f (CH_2Cl_2/MeOH 10:1) 0.70; ^{1}H NMR (300 MHz, $(CD_3)_2SO$): δ 8.30 (br s, 1H), 7.98 (s, 1H), 7.49 (s, 1H), 7.40 (d, 2H, J = 7.2 Hz), 7.35–7.20 (m, 7H), 6.89 (d, 4H, J = 8.5 Hz), 6.23 (t, 1H, J = 6.8 Hz), 6.04 (t, 1H, J = 6.5 Hz), 4.46 (m, 4H), 4.22 (m, 1H), 4.10 (m, 1H), 3.83 (m, 2H), 3.73 (s, 6H), 3.71 (dd, 1H, J = 3.7, 9.8 Hz), 3.65 (dd, 1H, J = 4.4, 10.9 Hz), 3.26 (d, 1H, J = 6.5 Hz), 3.20 (dd, 1H, J = 4.5, 10.9 Hz), 3.01 (s, 3H), 2.93 (s, 3H), 2.72 (m, 1H), 2.3–2.0 (m, 3H), 1.46 (s, 3H), 0.88 (s, 9H), 0.86 (s, 9H), 0.78 (s, 9H), 0.09 (s, 6H), 0.03 (s, 3H), 0.02 (s, 3H), 0.00 (s, 3H), −0.04 (s, 3H); ^{13}C NMR (75 MHz, $(CD_3)_2SO$): δ 162.54, 158.08, 157.17, 156.66, 150.22, 147.56, 144.37, 136.34, 135.12, 135.06, 133.62, 129.55, 127.73, 127.51, 126.69, 118.78, 113.08, 108.43, 86.85, 85.85, 85.22, 84.42, 82.14, 71.98, 71.26, 62.62, 54.91, 40.38, 39.81, 39.48, 39.09, 34.43, 25.59, 25.52, 25.43, 20.59, 17.80, 17.56, 17.41, 13.92, 12.37, -4.91; ESI MS (M + Na^+): 1258.3 (calcd 1258.76).

1-[N^1-(2′-deoxyguanosinyl)]-2-(N^3-thymidinyl)ethane **(14)**. At room temperature, tetrabutylammonium fluoride (167mg, 0.53 mmol) was added to a solution of 1 − {N^1-[3′,5′-*O*-(bis(*tert*-butyldimethylsilyl))-2-*N*-((dimethylamino)methylene)-2′-deoxyguanosinyl]}-2-{N^3-[5′-*O*-dimethoxytrityl-3′-*O*-(*tert*-butyldimethylsilyl)-thymidinyl]}ethane **13** (146 mg, 0.118 mmol) in THF (2 ml). After evaporation of the solvent and silica gel column chromatography (CH_2Cl_2:MeOH 15:1) the product was dissolved in MeOH (0.2 ml) and treated with concentrated solution of NH_4OH (200 µl). The reaction mixture was heated at 55°C for 17 h, then the solvent was evaporated and the crude product was dissolved in MeOH (320 µl) and water (80 µl) followed by acetic acid (1.6 ml). The reaction mixture was left at room temperature for 30 min after which the solvent was evaporated. The crude product was dissolved in H_2O and washed with ethyl acetate (3 × 2 mL). Evaporation of the water gave **14** (20 mg, 32%) as an oil. R_f (CH_2Cl_2/MeOH 5:1) 0.13; ^{1}H NMR (300 MHz, $(CD_3)_2SO$): δ 7.88 (s, 1H), 7.76 (s, 1H), 7.09 (br s, 2H), 6.12 (t, 1H, J = 6.5 Hz), 6.08 (t, 1H, J = 6.64 Hz), 5.24 (d, OH, J = 3.8 Hz), 5.20 (d, OH, J = 4.3 Hz), 5.03 (t, OH, J = 5.2 Hz), 4.90 (t, OH, J = 5.6 Hz), 4.33 (m, 1H), 4.22 (m, 2H), 4.15–4.04 (m, 3H), 3.78 (m, 2H), 3.62–3.49 (m, 4H), 2.49 (m, 1H), 2.20 (m, 1H), 2.1–2.0 (m, 2H), 1.75 (s, 3H); ^{13}C NMR (75 MHz, $(CD_3)_2SO$): δ 162.88, 156.47, 153.66, 150.41, 148.95, 135.18, 134.77, 115.53, 108.15, 87.40, 87.30, 84.84, 82.11, 70.61, 69.87, 61.66, 60.92, 39.62, 39.43, 38.38, 38.28, 12.72; ESI MS (M + Na^+): 558.2 (calcd 558.5).

N^3-(2-Hydroxyethyl-)-5′-O-dimethoxytrityl-3′-O-(tert-butyldimethylsilyl)thymidine **(15)**. At room temperature, NaOH (1M, 400 µl) was added to a solution of N^3-(2-iodoethyl-)-5′-*O*-dimethoxytrityl-3′-*O*-(*tert*-butyldimethylsilyl)thymidine **11** (297 mg, 0.36 mmol) in THF (2 ml). The reaction mixture was heated at 45°C for 17 h, cooled to room temperature, treated with a saturated solution of NH_4Cl, extracted with EtOAc. The organic layer was washed with brine and H_2O and dried over Na_2SO_4. Evaporation and silica gel column chromatography (hexane:EtOAc 2:1) gave **15** (180 mg, 71%) as a foam. R_f (Hexane/EtOAc 2:1) 0.38; ^{1}H NMR (300 MHz, $(CD_3)_2SO$): δ 7.59 (d, 1H, J = 1.1 Hz), 7.40 (dd, 2H, J = 7.13 Hz), 7.34–7.22 (m, 7H), 6.90 (d, 4H, J = 8.57 Hz), 6.21 (t, 1H, J = 6.4 Hz), 4.73 (t, OH, J = 5.8 Hz), 4.46 (m, 1H), 3.90 (d, 2H, J = 6.2 Hz), 3.84 (dd, 1H, J = 4.1, 7.2 Hz), 3.70 (s, 6H), 3.49 (dd, 2H, J = 6.5, 12.4 Hz), 3.29 (dd, 1H, J = 2.6, 10.6 Hz), 3.19 (dd, 1H, J = 4.3, 10.6 Hz), 2.32 (m, 1H), 2.18 (m, 1H), 1.55 (s, 3H), 0.79 (s, 9H), 0.01 (s, 3H), 0.00 (s, 3H); ^{13}C NMR (75 MHz, $(CD_3)_2SO$): δ 162.75, 158.09, 150.23, 144.41, 135.18, 135.08, 134.11, 129.57, 127.72, 127.53, 126.69, 113.09, 108.53, 85.83, 85.14, 84.55, 71.22, 62.71, 57.29, 54.91, 42.48, 39.41, 25.43, 17.43, 12.38, −4.49; ESI MS (M + Na^+): 725.4 (calcd 725.9).

1-{O^6-[3′,5′-O-(bis(tert-butyldimethylsilyl))-2′-deoxyguanosinyl]}-2-{N^3-[5′-O-(dimethoxytrityl)-3′-O-(tert-butyldimethylsilyl)thymidinyl]}ethane **(17)**. At 0°C, NaH (9.1 mg, 0.38 mmol) was added to a solution of N^3-(2-hydroxyethyl-)-5′-*O*-dimethoxytrityl-3′-*O*-(*tert*-butyldimethylsilyl)-thymidine **15** (65 mg, 0.12 mmol) and 2-amino-6-chloro-9-[2′-deoxy-3,5-bis-O-(*tert*-butyldimethylsi1y1)-β-D-ribofuranosyl]purine **16** (29,30) (91 mg, 0.12 mmol) in DME. The reaction mixture was left at room temperature for 1 h, treated with a saturated solution of NH_4Cl, extracted with EtOAc, washed with brine and H_2O and dried over Na_2SO_4. Evaporation and silica gel column chromatography (hexane/EtOAc 3:1) gave **17** (110 mg, 74%) as a white foam. R_f (Hexane/EtOAc 1:1) 0.58; ^{1}H NMR (500 MHz, $(CD_3)_2SO$): δ 8.01 (s, 1H), 7.62 (s, 1H), 7.38 (d, 2H, J = 7.4 Hz), 7.32–7.21 (m, 7H), 6.89 (d, 4H, J = 8.5 Hz), 6.34 (br s, 2H), 6.19 (t, 1H, J = 6.7 Hz), 6.14 (t, 1H, J = 6.5 Hz), 4.60 (t, 2H, J = 5.7 Hz), 4.50 (m, 1H), 4.45 (m, 1H), 4.25 (t, 2H, J = 5.5 Hz), 3.82 (m, 2H), 3.72 (s, 6H), 3.71 (d, 1H, J = 7.0 Hz), 3.63 (dd, 1H, J = 4.3, 10.8 Hz), 3.27 (dd, 1H, J = 1.9, 10.5 Hz), 3.16 (dd, 1H, J = 4.9, 10.5 Hz), 2.71 (m, 1H), 2.32–2.21 (m, 2H), 2.13 (m, 1H), 1.51 (s, 3H), 0.89 (s, 9H), 0.86 (s, 9H), 0.78 (s, 9H), 0.10 (s, 6H), 0.03 (s, 6H), 0.00 (s, 3H), −0.05 (s, 3H); ^{13}C NMR (125 MHz, $(CD_3)_2SO$): δ 162.59, 159.98, 159.61, 158.18, 154.04, 150.24, 144.48, 137.37, 135.26, 135.15, 134.41, 129.68, 129.64, 127.84, 127.62, 126.79, 113.89, 113.20, 108.58, 86.95, 85.94, 85.30, 84.81, 82.31, 72.20, 71.24, 62.73, 62.28, 59.68, 55.01, 39.43, 39.33, 38.68, 25.72, 25.64, 25.54, 20.70, 17.91, 17.66, 17.53, 14.03, 12.41, −4.82, −4.99, −5.17, −05.53, −5.57; ESI MS (M + Na^+): 1202.6 (calcd 1203.6).

1-[O^6-(2′-deoxyguanosinyl)]-2-(N^3-thymidinyl)ethane **(18)**. At room temperature, tetrabutylammonium fluoride (48 mg, 0.15 mmol) was added to a solution of 1-{O^6-[3′,5′-*O*-(bis(*tert*-butyldimethylsilyl))-2′-deoxyguanosinyl]}-2-{N^3-[5′-*O*-(dimethoxytrityl)-3′-*O*-(*tert*-butyldimethylsilyl)-thymidinyl]}ethane **17** (36 mg, 0.03 mmol) in THF (1 mL). The reaction mixture was left at room temperature for 2 h. After evaporation of the solvent and purification by silica gel column chromatography (EtOAc:MeOH 15:1) the product was dissolved in MeOH (320 µl) and water (80 µl) followed by acetic acid (1.6 ml). The reaction mixture was left at room

temperature for 3 h and the solvent was then evaporated. The crude product was dissolved in H_2O, washed with EtOAc (3 × 2 ml). Evaporation of the water gave **18** (14 mg, 85%) as a white foam. R_f (CH_2Cl_2/MeOH 5:1) 0.40; ^{1}H NMR (300 MHz, $(CD_3)_2SO$): δ 8.06 (s, 1H), 7.77 (s, 1H), 6.34 (br s, 2H), 6.19 (dd, 1H, J = 6.2, 7.6 Hz), 6.16 (t, 1H, J = 7.0 Hz), 5.26 (br s, OH, J = 3.8 Hz), 5.23 (br s, OH), 5.00 (br s, OH), 4.99 (t, OH, J = 4.9 Hz), 4.61 (m, 1H), 4.55 (m, 1H), 4.35 (m, 1H), 4.29 (dd, 1H, J = 5.5, 13 Hz), 4.25 (dd, 1H, J = 5.4, 11 Hz), 4.22 (m, 1H), 3.82 (m, 1H), 3.76 (m, 1H), 3.59 (m, 1H), 3.55 (m, 2H), 3.50 (m, 1H), 2.58 (m, 1H), 2.20 (m, 1H), 2.1–2.0 (m, 2H), 1.80 (s, 3H); ^{13}C NMR (75 MHz, $(CD_3)_2SO$): δ 162.65, 159.99, 159.50, 153.91, 150.43, 137.73, 134.89, 113.91, 108.41, 87.58, 87.37, 84.79, 82.74, 70.74, 70.21, 62.31, 61.72, 61.16, 39.55, 39.47, 39.34, 12.87; ESI MS (M + Na^+): 558.2 (calcd 558.5).

Oligonucleotides synthesis and purification

All oligonuceotides were synthesized on a PerSpective Biosystems Expedite 8909 DNA synthesizer on a 1 μmol scale using the standard ultra-mild protocol. Ultra-mild phosphoramidites and Q support were purchased from Glen Research. 6-carboxy fluorescein (6-FAM) phosphoramidite was purchased from Biosearch Technologies. All phosphoramidites were dissolved to a concentration of 0.1 M in anhydrous CH_3CN. Phenoxyacetic anhydride (0.25 M) in THF was used as a capping reagent. For the modified phosphoramidite the coupling time was extended to 15 min. The coupling yield of the modified phosphoramidite was >97%. The oligonucleotide **5** was removed from the solid support and deprotected by treatment with diisopropylamine/MeOH (1:10) for 17 h at room temperature, and purified by reverse phase HPLC using the following gradient: linear 0–10% B over 10 min, linear 10–20% B until 22 min, linear 20–25% B until 29 min, linear 25–35% B until 33 min, isocratic 100% B until 39 min, isocratic 0% B until 45 min; buffer A: 0.1 M TEAA (pH 7); and buffer B: CH_3CN. The peak of the oligonucleotide **5** eluted between 20.8 and 21.3 min. Nano-ESI MS: 5366.2 (calcd 5367.0).

Synthesis of cross-linked DNA

Oligonucleotide **5** (60 μM) was mixed with oligonucleotide **6** (12 μM) in 100 mM Hepes (pH 6.3), 300 mM NaCl, and 8 mM NaI. The mixture was left at room temperature in the dark for 12 days, and the cross-linked product was purified by JASCO HPLC using the following gradient: linear 0–20% B over 22 min, linear 20–25% B until 29 min, linear 25–35% B until 33 min, isocratic 100% B until 39 min, isocratic 0% B until 45 min; buffer A: 0.1 M TEAA (pH 7); and buffer B: CH_3CN. The peak of the cross-linked DNA eluted between 28 and 29.5 min. Nano-ESI MS: 10957.0 (calcd 10957.4).

Enzymatic digestion of cross-linked DNA

Interstrand cross-linked DNA (4 nmol) was incubated for 17 h at 30°C with 10 μl of snake venom phosphodiesterase (1 U) and exonuclease III (5 μl, 500 U) in a buffered solution containing 50 mM Tris–HCl, 25 mM $MgCl_2$, 1 mM $CaCl_2$, 96 mM NaCl at pH 8. The final volume was adjusted to 50 μl. The reaction mixture was then incubated for 10 min at 30°C with 0.5 μl (5 U) of calf intestine alkaline phosphatase (CIP), diluted with H_2O (150 μl) and directly injected into the HPLC using the following gradient: linear 0–10% B over 25 min, linear 10–35% B until 40 min, isocratic 100% B until 45 min, isocratic 0% B until 50 min; buffer A: 0.1 M TEAA (pH 7); and buffer B: CH_3CN. The resulting peaks were identified by co-injection with the corresponding standards and eluted at the following times: dC (6.1 min), dG (13.0 min), T (14.3 min), dA (20.2 min), **14** (35.6 min), cross-link dimer **18** (37.2 min), 6-FAM (41.7 min).

RESULTS AND DISCUSSION

Design of a sequence-specific alkylation reaction for the generation of DNA interstrand crosslinks

We set out to design a strategy that would allow for the incorporation of a precursor of an ethylene-bridged ICL into one strand of DNA using solid phase synthesis and specific alkylation reaction to a complementary strand following annealing. In our design we first considered the reaction of the CENUs with DNA itself. CENUs initially react with the oxygen of guanine base to form O^6-chloroethylguanine, which then undergoes intramolecular cyclization to give rise to N^1,O^6-ethanoguanine (Figure 1A). This positively charged intermediate is highly reactive and can undergo nucleophilic attack by the N-3 of cytosine located on the opposite strand of DNA (14). We reasoned that although the O^6-chloroethylguanine or N^1,O^6-ethanoguanine intermediates would likely be too unstable to be incorporated into DNA by solid phase synthesis (31), it might be possible to mimic this process in an alkylation reaction with a different reactive precursor in one strand. In a related study at the (mono) nucleoside level, Miller and coworkers synthesized a thymidine dimer crosslinked at their N-3 positions through an ethylene bridge using a 3-iodoethylthymidine derivative as a precursor, which they subsequently introduced into DNA as a dimer (20).

We reasoned that it might be possible to use a similar reaction in the context of dsDNA. Preliminary studies using 1-(2-chloroethyl)-guanosine and 3-(2-chloroethyl)-cytosine (Figure 1B) revealed that these molecules undergo rapid cyclization to the N^1,N^2-ethanoguanine and N^3,N^4-ethanocytosine, respectively, and would be unsuitable precursors for a C-G crosslink. In contrast, it appeared that the corresponding 3-(2-chloroethyl)thymidine would be a suitable ICL precursor with a stability profile that might allow its incorporation into DNA using solid phase synthesis (see below). We therefore set out to incorporate 3-(2-chloroethyl)thymidine into DNA and test its ability to mediate site-specific ICL formation.

Preparation of N-3-chloroethylthymidine and its incorporation into DNA

The synthesis of phosphoramidite **4** containing the chloroethyl thymidine moiety was accomplished as shown in Figure 2. Alkylation of 5′-*O*-dimethoxytrityl-3′-*O*-*tert*-butyldimethylsilylthymidine **1** (28) at N-3 position using Cs_2CO_3 and bromochloroethane gave N-3-chloroethyl product **2** (32). The silyl group was removed from **2** using

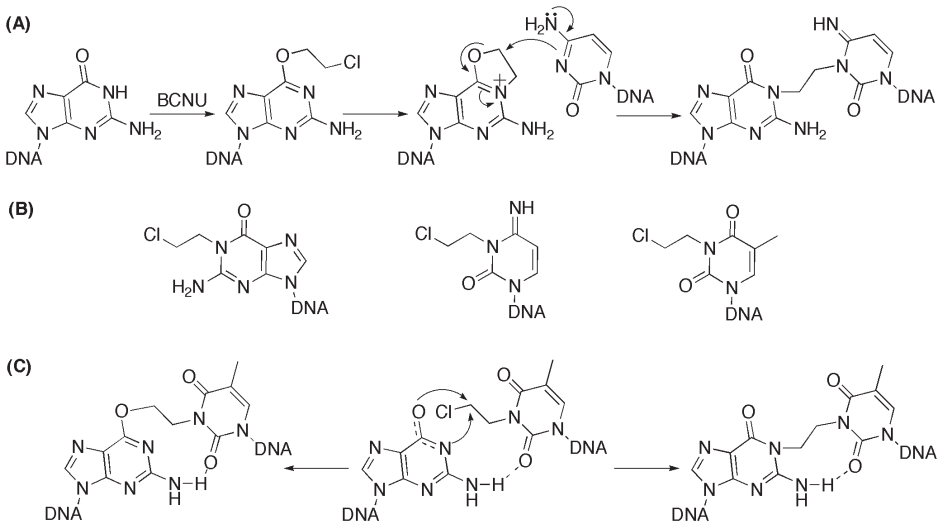

Figure 1. Formation of CENU and CENU-like interstrand crosslinks. (**A**) Mechanism of ICL formation by CENUs. (**B**) Reactive G, C and T derivatives considered as cross-linking precursors in this study (**C**) Cross-linking reaction using 3-(2-chloroethyl)thymidine moiety in DNA may proceed through the O-6 or N-1 site of guanine.

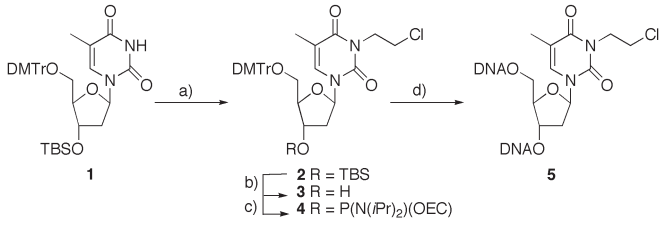

Figure 2. Synthesis of oligonucleotides containing N-3-chloroethylthymidine. Reagents and conditions: (a) Cs_2CO_3, $BrCH_2CH_2Cl$, DMF, 86%. (b) TBAF, THF, 89%. (c) $ClP(N(iPr)_2)(OEC)$, $(iPr)_2NEt$, CH_2Cl_2, 90%. (d) Solid-phase DNA synthesis and deprotection.

tetrabutylammonium fluoride to give alcohol **3**, which was converted to the phosphoramidite **4**. The building block **4** was thus obtained in a straightforward manner in 68% overall yield from **1**.

Since we were concerned whether the chloroethyl group would be stable toward base induced hydrolysis, we tested its stability toward very mild basic deprotection conditions that we had previously developed (33). Chloroethyl **3** proved to be stable toward hydrolysis upon extended treatment in diisopropylamine/methanol (1:10) (data not shown). We therefore proceeded to incorporate phosphoramidite **4** into oligonucleotides using 'ultramild' phosphoramidites (*i*PrPac-dG, Pac-dA and Ac-dC) and Q support (34) and subsequently removed the oligonucleotide from solid support and removed base and phosphodiester bond-protecting groups by treatment with a 10% solution of diisopropylamine in methanol at room temperature for 17 h. The main product was purified by HPLC and analyzed by nano-ESI to confirm its identity and purity. The observed molecular weight (5366.2) was indeed consistent with the calculated molecular weight (5367.0) for oligonucleotide **5**.

Site-specific ICL formation

With the chloroethyl thymine-containing oligonucleotide **5** in hand, we tested its ability to form interstrand crosslinks with complementary oligonucleotides. We annealed **5** to a complementary oligonucleotide **6** containing a 6-carboxyfluorescein (6-FAM) at the 5′-end and a G opposite to the reactive group. Following incubation of the two oligonucleotides at room temperature, we observed the formation of a band with slower migration in a 20% denaturing PAGE gel, indicative of ICL formation. The band increased over the course of 2 weeks, eventually yielding ~20% of the FAM-labeled oligonucleotide in crosslinked form (Figure 3B, lanes 2–4). The reaction was of similar efficiency in a variety of buffers (ammonium acetate, sodium acetate, Hepes·KOH, Tris–HCl, ammonium chloride) and over a wide pH range (5–12) (data not shown). We carried out our studies in Hepes·KOH at pH 6, which appeared to afford the cleanest reaction. Although chloride **5** was stable in aqueous solution at room temperature for several days, after 15 days we observed an increasing amount of the corresponding alcohol by mass spectrometry (data not shown). We then studied the influence of the concentrations of oligonucleotides **5** and **6** on the efficiency of the cross-linking reaction. We observed an increase in reaction efficiency up to about five equivalents of oligoncleotide **5** over **6** (Figure 3B, lanes 5–9). Under these conditions, we consistently achieved a cross-linking yield of 25–30% based on the unmodified strand (Figures 3B and 4). We then compared the ability of **5** to react with the four DNA bases. Oligonucleotide **5** exhibited the highest yield of crosslink formation if paired with guanine with the reactivity decreasing in the order G ≫ A > C ≫ T (Figure 3B, lanes 10–13).

Characterization of cross-linked oligonucleotide 10

The main crosslinked product between oligonucleotides **5** and **6** was purified by HPLC (Figure 4B) and its mass determined by nano-ESI-MS. The observed molecular weight (10957.0) was found to be consistent with the calculated mass (10957.4) of the cross-linked product. Since two products of identical mass can be formed by reaction of the chloride in **5** with either the N-1 or O-6 position of the guanine base (Figure 1C), we wished to determine the structure of the

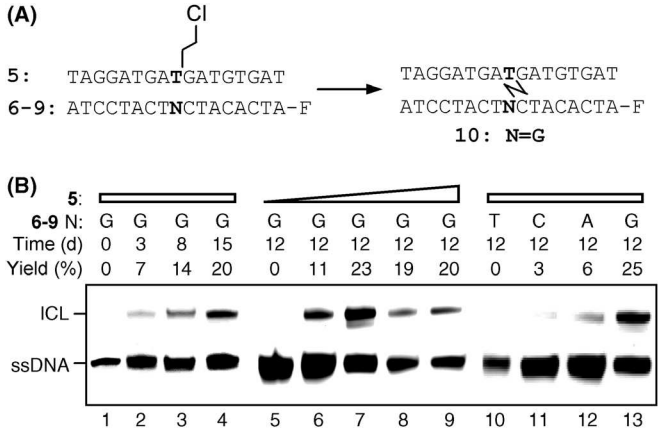

Figure 3. (**A**) Schematic representation of the cross-linking reaction; 3-(2-chloroethyl)-T containing oligonucleotide **5** was annealed to complementary fluorescently labeled oligonucleotides **6–9** (**6**: N = G; **7**: N = A; **8**: N = C; **9**: N = T; F = 6-carboxy fluorescein, 6-FAM) and ICL formation was analyzed by 20% denaturing PAGE. (**B**) Analysis of ICL formation by 20% denaturing PAGE. The oligonucleotides used, incubation times, and yields are indicated above the gel. The following concentrations of substrates were used: oligonucleotide **5**: lanes 1–4, 36 mM; lane 5, 12 mM; lane 6, 24 mM; 7, 10–13, 60 mM; lane 8, 84 mM; lane 9, 108 mM; oligonucleotides **6–9**: 12mM.

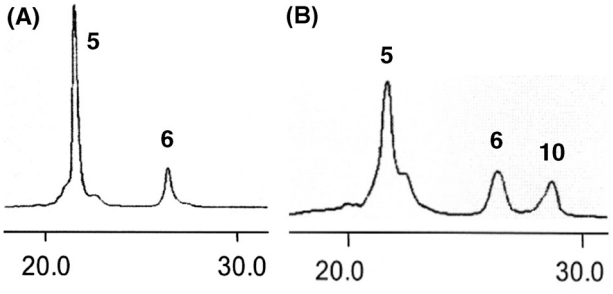

Figure 4. HPLC analysis of T-G interstrand crosslinks. (**A**) HPLC profile of the oligonucleotides **5** and **6** before the cross-linking reaction. (**B**) HPLC profile of formation of ICL **10** after incubation of **5** and **6** for 12 days at room temperature.

crosslink by nucleoside composition analysis and comparison of the digested product with an authentic standard. We therefore needed to synthesize the cross-linked dimers **14** and **18**, containing dG alkylated at N-1 and O-6, respectively (Figure 5). Initial attempts to alkylate the N-1 position of the protected-2′-deoxyguanosine **12** (32) with chloroethyl compound **2** using different alkylating conditions were not successful and only the starting materials were isolated. We reasoned that improving the reactivity of the alkylating agent might lead to success and replaced the chloride in **2** with iodide, using NaI in acetonitrile. Silver(I) promoted the alkylation of iodide **11** to yield the fully protected dimer **13**. By contrast, the use of basic conditions (Cs_2CO_3 or DPU in organic solvents) did not lead to alkylation and only yielded the corresponding alcohol **15**. The fully deprotected N1-linked dimer **14** was obtained by consecutive desilylation using tetrabutylammonium fluoride, deprotection of the exocyclic amine using NH_4OH, and removal of the trityl group using 80% acetic acid.

The fully protected O6-linked dimer **17** was obtained by reaction of hydroxyethyl thymine **15**, obtained by basic hydrolysis of iodide **11**, and protected 6-chloropurine **16** (29,30) in DME using NaH as a base (Figure 5). Other attempts, using DMF as a solvent gave a complex mixture of products, whereas a Mitsunobu reaction between hydroxyethyl **15** and 3′,5′-*O*-(bis(*tert*-butyldimethylsilyl))-2′-deoxyguanosine yielded no detectable product. The fully deprotected O6-linked dimer **18** was obtained by desilylation followed by acid hydrolysis of the trityl group.

Although the two dimers **14** and **18** have identical signals for their molecular weight ions in MS analysis, they can be distinguished by ^{13}C-NMR spectroscopy based on their signals for the methylene groups. The ^{13}C-NMR spectrum of dimer **14** shows signals at 61.66 and 60.92 p.p.m., corresponding to the two O-CH_2 groups of the C-5′ deoxyriboses, and signals at 39.62, 39.43, 38.38, 38.28 p.p.m., corresponding to N-CH_2CH_2-N and C-2′ of the two deoxyribose residues, respectively. The ^{13}C-NMR spectrum of compound **18** showed signals at 62.31, 61.72, 61.16 p.p.m., indicating the presence of three O-CH_2 groups: two for C-5′ of the deoxyriboses, and one for the O-ethylene bridge; the other three signals, at 39.55, 39.47, 33.34 p.p.m., correspond to N-CH_2 and the C-2′ of the two deoxyribose moieties.

The availability of the authentic standards **14** and **18** allowed us to identify the structure of the ICL in oligonucleotide **10**. Cross-link oligonucleotide **10** was enzymatically digested, dephosphorylated and analyzed by HPLC (Figure 6B). In addition to the four natural nucleosides dC, dG, T and dA, we observed two additional peaks; the first one corresponded to the fluorescent label at the 5′ end of the G containing oligonucleotide. The second one co-eluted with the O6-linked dimer **18** with a retention time of 37.2 minutes, clearly distinct from **14**, which eluted at 35.6 min. The nucleoside composition analysis therefore establishes that the crosslink is formed exclusively between N-3 of T and O-6 of dG.

Mechanism of ICL formation between N-3 of T and O-6 of G

The exclusive reaction of the chloroethyl group of T with O-6 of dG and absence of reaction with N-1 of dG deserves some comment. It appears that the chlorine in 3-(2-chloroethyl)thymidine is very inert due to the two electron-withdrawing groups in the pyrimidine ring, rendering it unreactive toward direct substitution by oxygen or nitrogen nucleophiles. This is illustrated by the extremely low reactivity of **2** toward nucleophilic substitution under aprotic conditions (Figure 5). Likewise, a direct S_N2 substitution of the chloride in the context of dsDNA is not a likely event. However, under solvolytic conditions, the formation of the ionic intermediate **19** (Figure 7) may occur at low frequency due to the intramolecular nature of this cyclization reaction. The ionic character of **19** causes it to react exclusively with the oxygen of the amide bond in dG, due to the latter's partial anionic character. This mechanism accounts for the very slow ICL formation, because the frequency of the formation of **19** is very low. Once it is formed it is likely to react very quickly with the closest nucleophile with anionic character, in this case O-6 of dG.

This preferential reactivity of oxygen-centered rather than nitrogen-centered nucleophiles with electrophiles with ionic

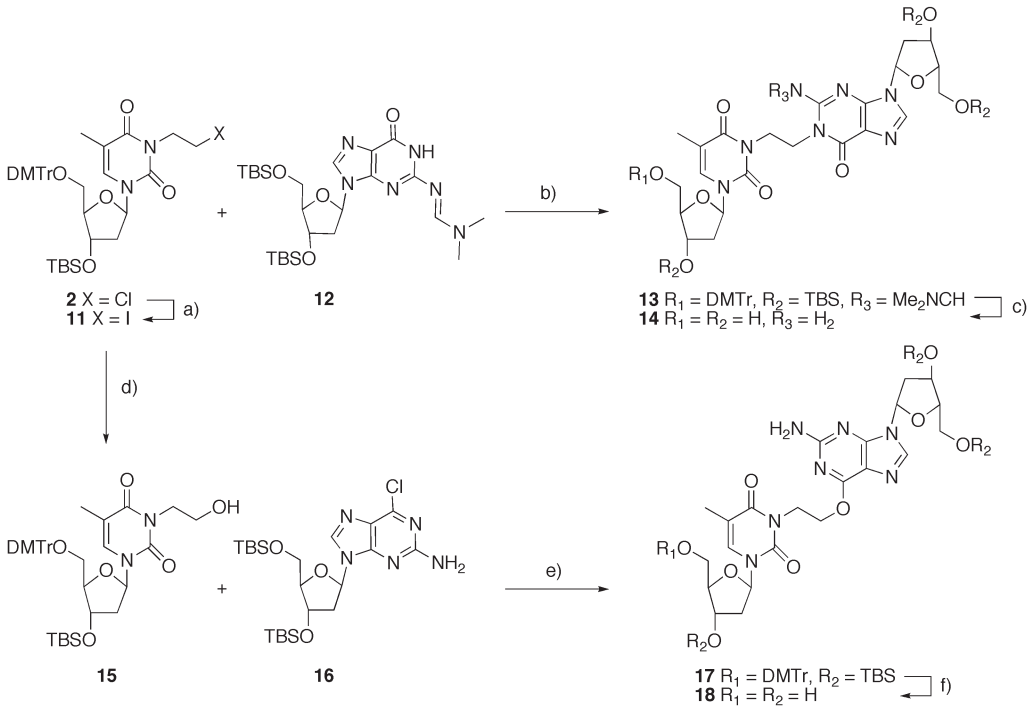

Figure 5. Synthesis of crosslinked dimers 14 and 18. Reagents and conditions: a) NaI, CH_3CN, 70°, 86%. b) Ag_2O, DMF, 48%. c) i) TBAF, THF. ii) NH_4OH. iii) 80% CH_3COOH, 32%. d) NaOH, THF, 71%. e) NaH, DME, 74%. f) i) TBAF, THF. ii) 80% CH_3COOH, 85%.

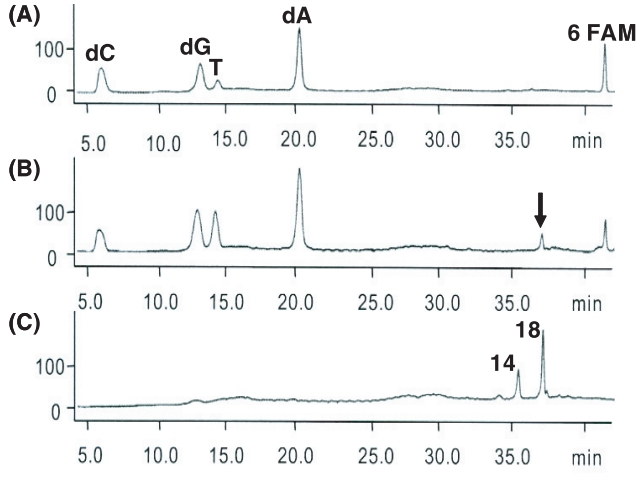

Figure 6. Enzymatic digesion analysis of an oligonucleotide containing a G-T ICL. (**A**) HPLC profile of oligonucleotide **6** following enzymatic digestion and dephosphporylation. The four nucleosides and fluorescent label 6-FAM are indicated. (**B**) HPLC profile of ICL-containing oligonucleotide **10** after enzymatic digestion and dephosphorylation. The peak for the crosslinked T-G residue is denoted by an arrow (**C**) HPLC profiles of the T(N3)-dG(N1) dimer **14** and the T(N3)-dG(O6) dimer **18**.

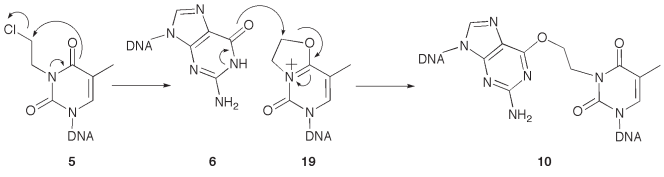

Figure 7. Proposed mechanism for the formation of the (O6)G-(N3)T crosslink. An intramolecular substitution of the chloride in **5** leads to the positively charged N^3,O^4-ethanothymidine (19). Equivalent formation of O^2,N^3-ethanothymidine might also occur. **19** reacts exclusively with O-6 of the opposite guanine base due to more anionic character of O-6 versus N-1 of G.

character is also observed in the reaction of alkylating agents with the DNA bases. Alkylating agents such as N-methyl-N-nitrosourea (NMU), which react with an S_N1-like mechanism with cationic character in the transition state, are highly oxophilic and show high reactivity with O-6 of dG (35,36). Less reactive alkylating agents that react by an S_N2-type mechanisms without significant positive charge in the transition state instead prefer to react with nitrogen-based nucleophile in DNA such as N-7 of G.

Our results suggest that the development of a related strategy for the preparation of the (N-1)G–(N-3)C crosslink, formed by various CENUs would likely have to rely on a precursor that allows the *in situ* formation of the charged N^1,O^6-ethanoguanosine intermediate. In this case, where a C base is situated to act as a nucleophile to react with the ethanoguanosine cation, the N-3 of C, which contains more anionic character than N-4, is known to react as the nucleophile (14,27). In the interim the G-T crosslink generated here will serve as a useful mimic for the study of ethylene-bridged ICLs for the investigation of the repair pathways for this type of lesion.

CONCLUSIONS

We have developed a simple and concise approach for the synthesis of an ethylene bridged DNA interstrand crosslink. For this purpose we have synthesized an oligonucleotide carrying a chloroethyl moiety at a defined T base poised to

react selectively with a base on a complementary DNA strand. To our knowledge this constitutes the first example of an oligonucleotide with an alkyl chloride side chain and one of a few examples of a specific ICL formation using a reactive group on one strand of DNA (21–23). Since oligonucleotides containing 2-chloroethylthymidine have in principle the ability to react with different types of DNA in a sequence-specific manner, they might also be generally useful for the sequence-specific covalent modification of DNA. Given the slow kinetics of the cross-linking reaction, it remains to be seen whether it will be useful in the context of gene-modification using antisense or triple-helix approaches (37–39).

SUPPLEMENTARY DATA

Supplementary data are available at *NAR* online.

ACKNOWLEDGEMENTS

The authors are grateful to The Vinh Ho, Angelo Guainazzi and in particular Prof. Francis Johnson for comments on the manuscript. This work was supported by the European Commission (LSHB-CT-2004-005204), the Roche Research Foundation (126–2003), the Swiss National Science Foundation (3130-054873.98) and the New York State Office of Science and Technology and Academic Research (NYSTAR) grant No. C040069. Funding to pay the Open Access publication charges for this article was provided by was provided by the European Commission (LSHB-CT-2004-005204).

Conflict of interest statement. None declared.

REFERENCES

1. Noll,D.M., Mason,T.M. and Miller,P.S. (2006) Formation and repair of interstrand cross-links in DNA. *Chem. Rev*., **106**, 277–301.
2. Schärer,O.D. (2005) DNA interstrand crosslinks: natural and drug-induced DNA adducts that induce unique cellular responses. *ChemBioChem*., **6**, 27–32.
3. Joenje,H. and Patel,K.J. (2001) The emerging genetic and molecular basis of Fanconi anaemia. *Nat. Rev. Genet*., **2**, 446–457.
4. D'Andrea,A.D. and Grompe,M. (2003) The Fanconi anaemia/BRCA pathway. *Nat. Rev. Cancer*, **3**, 23–34.
5. McHugh,P.J., Spanswick,V.J. and Hartley,J.A. (2001) Repair of DNA interstrand crosslinks: molecular mechanisms and clinical relevance. *Lancet Oncol*., **2**, 483–490.
6. Panasci,L., Paiement,J.P., Christodoulopoulos,G., Belenkov,A., Malapetsa,A. and Aloyz,R. (2001) Chlorambucil drug resistance in chronic lymphocytic leukemia: the emerging role of DNA repair. *Clin. Cancer Res*., **7**, 454–461.
7. Schärer,O.D. (2003) Chemistry and biology of DNA repair. *Angew. Chem. Int. Ed. Engl*., **42**, 2946–2974.
8. Friedberg,E.C., Walker,G.C., Siede,W., Wood,R.D., Schultz,R.A. and Ellenberger,T. (2005) *DNA Repair and Mutagenesis*. 2nd edn. ASM Press, Washington DC.
9. Dronkert,M.L. and Kanaar,R. (2001) Repair of DNA interstrand cross-links. *Mutat. Res*., **486**, 217–247.
10. Noll,D.M., Noronha,A.M., Wilds,C.J. and Miller,P.S. (2004) Preparation of interstrand cross-linked DNA oligonucleotide duplexes. *Front. Biosci*., **9**, 421–437.
11. Ojwang,J.O., Grueneberg,D.A. and Loechler,E.L. (1989) Synthesis of a duplex oligonucleotide containing a nitrogen mustard interstrand DNA–DNA cross-link. *Cancer Res*., **49**, 6529–6537.
12. Rink,S.M., Solomon,M.S., Taylor,M.J., Rjur,S., McLaughlin,L.W. and Hopkins,P.B. (1993) Covalent structure of a nitrogen mustard-induced DNA interstrand cross-link: an N7-to-N7-linkage of deoxyguanosine residues at the duplex sequence 5′-d(GNC). *J. Am. Chem. Soc*., **115**, 2551–2557.
13. Warren,A.J. and Hamilton,J.W. (1996) Synthesis and structural characterization of the N2G-mitomycin C-N2G interstrand cross-link in a model synthetic 23 base pair oligonucleotide DNA duplex. *Chem. Res. Toxicol*., **9**, 1063–1071.
14. Fischhaber,P.L., Gall,A.S., Duncan,J.A. and Hopkins,P.B. (1999) Direct demonstration in synthetic oligonucleotides that N,N′-bis(2-chloroethyl)-nitrosourea cross links N1 of deoxyguanosine to N3 of deoxycytidine on opposite strands of duplex DNA. *Cancer Res*., **59**, 4363–4368.
15. Bhat,B., Leonard,N.J., Robinson,H. and Wang,A.H.-J. (1986) Deoxyadenosine and thymidine bases held proximal and distal by means of a covalently-linked dimensional analogue of dA·dT: intramolecular vs intermolcular hydrogen bonding. *J. Amer. Chem. Soc*., **118**, 10744–10751.
16. Harwood,E.A., Sigurdsson,S.T., Edfeldt,N.B.F., Reid,B.R. and Hopkins,P.B. (1999) Chemical synthesis and preliminary structural characterization of a nitrous acid interstrand cross-linked duplex DNA. *J. Am. Chem. Soc*., **121**, 5081–5082.
17. Li,H.Y., Qiu,Y.L., Moyroud,E. and Kishi,Y. (2001) Synthesis of dna oligomers possessing a covalently cross-linked Watson–Crick base pair model. *Angew. Chem. Int. Ed. Engl*., **40**, 1471–1475.
18. Noll,D.M., Noronha,A.M. and Miller,P.S. (2001) Synthesis and characterization of DNA duplexes containing an N(4)C-ethyl-N(4)C interstrand cross-link. *J. Am. Chem. Soc*., **123**, 3405–3411.
19. Noronha,A.M., Noll,D.M., Wilds,C.J. and Miller,P.S. (2002) N(4)C-ethyl-N(4)C cross-linked DNA: synthesis and characterization of duplexes with interstrand cross-links of different orientations. *Biochemistry*, **41**, 760–771.
20. Wilds,C.J., Noronha,A.M., Robidoux,S. and Miller,P.S. (2004) Mispair-aligned N^3T-alkyl-N^3T interstrand cross-linked DNA: synthesis and characterization of duplexes with interstrand cross-links of variable length. *J. Am. Chem. Soc*., **126**, 9257–9265.
21. Kobertz,W.R. and Essigmann,J.M. (1997) Solid-phase synthesis of oligonucleotides containing a site-specific psoralen derivative. *J. Am. Chem. Soc*., **119**, 5960–5961.
22. Hong,I.S. and Greenberg,M.M. (2005) DNA interstrand cross-link formation initiated by reaction between singlet oxygen and a modified nucleotide. *J. Am. Chem. Soc*., **127**, 10510–10511.
23. Hong,I.S. and Greenberg,M.M. (2005) Efficient DNA interstrand cross-link formation from a nucleotide radical. *J. Am. Chem. Soc*., **127**, 3692–3693.
24. Ferentz,A.E., Keating,T.E. and Verdine,G.L. (1993) Synthesis and characterization of disulfide cross-linked oligonucleotides. *J. Am. Chem. Soc*., **115**, 9006–9014.
25. Erlanson,D.A., Chen,L. and Verdine,G.L. (1993) DNA methylation through a locally unpaired intermediate. *J. Am. Chem. Soc*., **115**, 12583–12584.
26. Dooley,P.A., Tsarouhtsis,D., Korbel,G.A., Nechev,L.V., Shearer,J., Zegar,I.S., Harris,C.M., Stone,M.P. and Harris,T.M. (2001) Structural studies of an oligodeoxynucleotide containing a trimethylene interstrand cross-link in a 5′-(CpG) motif: model of a malondialdehyde cross-link. *J. Am. Chem. Soc*., **123**, 1730–1739.
27. Ludlum,D.B. (1997) The chloroethylnitrosoureas: sensitivity and resistance to cancer chemotherapy at the molecular level. *Cancer Invest*., **15**, 588–598.
28. Sung,W.L. (1981) Synthesis of 4-triazolopyrimidinone nucleotide and its application in synthesis of 5-methylcytosine-containing oligodeoxyribonucleotides. *Nucleic Acids Res*., **9**, 6139–6151.
29. Janeba,Z., Francom,P. and Robins,M.J. (2003) Efficient syntheses of 2-chloro-2′-deoxyadenosine (cladribine) from 2′-deoxyguanosine(1). *J. Org. Chem*., **68**, 989–992.
30. Nair,V. and Purdy,D.F. (1991) Synthetic approaches to new doubly modified nucleosides: congeners of cordycepin and related 2-deoxyadenosine. *Tetrahedron*, **47**, 365–382.
31. Noll,D.M. and Clarke,N.D. (2001) Covalent capture of a human O(6)-alkylguanine alkyltransferase-DNA complex using N(1),O(6)-ethanoxanthosine, a mechanism-based crosslinker. *Nucleic Acids Res*., **29**, 4025–4034.

32. Huang,Y., Torres,M.C., Iden,C.R. and Johnson,F. (2003) Regioselective synthesis of 1,N(2)-etheno-2′-deoxyguanosine and its generation in oligomeric DNA. *Chem. Res. Toxicol.*, **16**, 708–714.
33. Gillet,L.C., Alzeer,J. and Schärer,O.D. (2005) Site-specific incorporation of N-(deoxyguanosin-8-yl)-2-acetylaminofluorene (dG-AAF) into oligonucleotides using modified 'ultra-mild' DNA synthesis. *Nucleic Acids Res.*, **33**, 1961–1969.
34. Pon,R.T. and Yu,S. (1997) Hydroquinone-*O,O'*-diacetic acid as a more labile replacement of succinic acid linkers in solid-phase oligonucleotide synthesis. *Tetrahedron Let.*, **38**, 3327–3330.
35. Lawley,P.D. (1974) Some chemical aspects of dose–response relationships in alkylation mutagenesis. *Mutat. Res.*, **23**, 283–295.
36. Loechler,E.L. (1994) A violation of the Swain-Scott principle, and not SN1 versus SN2 reaction mechanisms, explains why carcinogenic alkylating agents can form different proportions of adducts at oxygen versus nitrogen in DNA. *Chem. Res. Toxicol.*, **7**, 277–280.
37. Sasaki,S. (2001) Active oligonucleotides incorporating alkylating an agent as potential sequence- and base selective modifier of gene expression. *Eur. J. Pharm. Sci.*, **13**, 43–51.
38. Vasquez,K.M. and Glazer,P.M. (2002) Triplex-forming oligonucleotides: principles and applications. *Q. Rev. Biophys.*, **35**, 89–107.
39. Ali,M.M., Oishi,M., Nagatsugi,F., Mori,K., Nagasaki,Y., Kataoka,K. and Sasaki,S. (2006) Intracellular inducible alkylation system that exhibits antisense effects with greater potency and selectivity than the natural oligonucleotide. *Angew. Chem. Int. Ed. Engl.*, **45**, 3136–3140.